Infant speech perception is known to undergo critical changes before the first year of life. Young infants are able to distinguish non-native sound contrasts, whereas older infants and adults lose this ability by attaining the native language (Werker & Tao 1994/2002; Kuhl et al. 2005; Alfano et al. 2006). This dramatic change in non-native speech perception has been shown to predict later language development (Kuhl et al. 2005; Rivera-Gaxiola et al. 2005).

Much of this research was conducted using the Conditioned Head Turn (CHT) procedure (Werker et al. 1997), which is a visually reinforced infant speech discrimination paradigm that assesses the discrimination of sound categories on the presentation of multiple exemplars. For adults, AB or ABD discrimination tasks are used (e.g., Tao et al. 2006).

Unlike the head-turning dependent measure, eye-movements are very fast, have low metabolic cost, and are automatically detected by eye-tracking systems. Furthermore, eye-movements can be used as a dependant measure in both infant and adult studies. Anticipatory eye-movements have been shown to cue the linguistic processing of sounds (Altmann & Kamide 2007; Kukona et al. 2011), and to measure the categorization of visual and auditory stimuli, as well as the learning of audio-visual contingencies by infants as young as 6 months (McMurty & Aslin 2004; Shukla et al. 2011; Bjerva et al. 2011).

### Experiment 1: ABX task

**Method**

Participants: 18 native Hindi speakers (6 females) 18 native EP speakers (12 females)

Stimuli: 2 natural speech pairs of tokens of the Portuguese native contrast /pa/- /ka/, produced by a female speaker, were used in the training phase.

Procedure: 3 pairs of tokens of the Portuguese contrast /pa/- /ka/ were used in the training phase, which consisted of 15 trials (with response feedback).

Results and discussion: Portuguese-speaking adults scored lower than Hindi speaking adults on Hindi /-a/- contrast discrimination, as well as their larger error rate (20.1% vs. 20.3%). A one-way ANOVA revealed a significant difference between the two groups (F(1,35) = 6.26, p < .05).

### Experiment 2 and 3: AEM paradigm

**Exp. 2: Method**

Participants: 9 native EP speakers (M=30; range 21-38; 5 females) Stimuli: The Hindi contrast /-a/- /-e/ in the test phase.

Procedure: Training phase: 3 blocks of 4 trials each. Test phase: 1 trial with an attention getter at the center of the screen (trigger ACS, 400 ms), followed by a dynamic visual stimulus at the top left side consistent with the presentation of the speech sound file. VAOI onset was delayed relative to the onset of the sound (200 ms). Two different re-inforcers were used: Text phase: 24 trials (12 no change trials: /-a/-/-e/ only, 12 change trials: /-a/-/-e/ or /-i/-/-e/). In the change trials, VAOI appears after 400 ms when the test sound to the correct side while side 2 is playing (fig. 3). Two reinforcers different from those in the training were used in the test phase.

**Exp. 3: Results and discussion**

- Infants showed no AEM in this training (a repeated measures ANOVA revealed no effect of Training block (F(2,18) = 1.43, p = .27; η² = .14). Non-parametric Wilcoxon signed ranks revealed no significant differences.
- Only 1 out of 9 show AEM in block 2 and in block 3.
- Moreover, an effect in the right direction emerged: Infants are looking quicker to VAOI in block 2 than in block 1, and in block 3 than in block 2. Possibly, the amount of training was not enough to trigger anticipatory looking (18 to 30 trials in previous studies).

### Discussion

- We developed an eye-tracking version of CHT by exploring anticipatory eye movement triggered by auditory stimuli to assess speech discrimination by adults and infants.
- Infants were found to present AEM after 12 training trials, and failed to discriminate a non-native speech sound contrast in the test phase. This result replicates findings from an ABD discrimination task, supporting the effect of language experience on adult discrimination of non-native contrasts.
- Our results from Exp.2 thus suggest that the AEM paradigm developed is successful to assess discrimination of speech sound categories.
- Infants did not present AEM, although there is evidence that they were learning to predict image appearance from the sound heard. Different factors may have affected infants' behavior: (i) amount of trials (12 training trials were not enough). (ii) attractiveness of the visual stimulus that hints infants' attention during the sound presentation; (iii) diversity of the visual stimuli that show different visual features were used.

### Conclusion

This study is the first to provide a fully-relevant and effective auditory-visual stimulus to assess adult and infant speech discrimination in an AEM paradigm. The initial results suggest that AEM is a promising tool for the assessment of speech perception in both adults and infants.